No. data points : n

No. labels : L

Micro averaging F-measure

$$\hat{TP}_{micro}(f) := \frac{1}{nL}\sum_{i=1}^{n}\sum_{j=1}^{L}\mathbb{I}\{f_j(x_i)=1, y_j^i=1\}$$

$$\hat{TN}_{micro}(f) := \frac{1}{nL}\sum_{i=1}^{n}\sum_{j=1}^{L}\mathbb{I}\{f_j(x_i)=0, y_j^i=0\}$$

$$\hat{\Theta}_{micro} := \frac{\sum_{i=1}^{n}\sum_{j=1}^{L}\mathbb{I}\{y_j^i=0\}}{\sum_{i=1}^{n}\sum_{j=1}^{L}\mathbb{I}\{y_j^i=1\}}$$

$$\hat{F}_{\beta}^{micro}(f) := \frac{(1+\beta^2)\cdot\hat{TP}_{micro}(f)}{\beta^2+\hat{\Theta}_{micro}+\hat{TP}_{micro}(f)-\hat{\Theta}_{micro}\cdot\hat{TN}_{micro}(f)}$$

$$\hat{F}_{\beta}^{micro}(f)\geq v \iff (1+\beta^2-v)\cdot\hat{TP}_{micro}(f)+v\cdot\hat{TN}_{micro}(f)\geq v(\beta^2+\hat{\Theta}_{micro})$$

AMP micro

$$v_{t+1} \longleftarrow \hat{F}_{\beta}^{micro}(f_t)$$

$$f_{t+1} \longleftarrow \operatorname*{argmax}_{f\in\mathcal{F}}\left\{\underbrace{(1+\beta^2-v_{t+1})}_{a_{t+1}}\cdot\hat{TP}_{micro}(f)+\underbrace{v_{t+1}\cdot\hat{\Theta}_{micro}}_{b_{t+1}}\hat{TN}_{micro}(f)\right\}$$

$$a_{t+1}\hat{TP}_{micro}(f)+b_{t+1}\hat{TN}_{micro}(f)=\frac{1}{nL}\sum_{i=1}^{n}\sum_{j=1}^{L}\left(a_{t+1}\hat{TP}_j^i(f)+b_{t+1}\hat{TN}_j^i(f)\right)$$

$$\boxed{\begin{aligned}a_{t+1}&=1+\beta^2-v_{t+1}\\ b_{t+1}&=v_{t+1}\cdot\hat{\Theta}_{micro}\end{aligned}}$$

Define
$$\begin{cases}\hat{TP}_j^i(f)=\mathbb{I}\{f_j(x_i)=1, y_j^i=1\}\\ \hat{TN}_j^i(f)=\mathbb{I}\{f_j(x_i)=0, y_j^i=0\}\end{cases}$$

$$=\frac{1}{L}\sum_{j=1}^{L}\frac{1}{n}\sum_{i=1}^{n}a_{t+1}\hat{TP}_j^i(f_j)+b_{t+1}\hat{TN}_j^i(f_j)$$

$$f_{t+1}=\{f_{t+1}^1, f_{t+1}^2, f_{t+1}^3, \ldots, f_{t+1}^L\}$$

f_{t+1}^j is obtained by solving a weighted binary classification problem on the j^{th} label points with the costs a_{t+1} and b_{t+1} on positives and negatives respectively.

Macro averaging F-measure

$$\hat{TP}^{j}_{macro}(f) = \frac{1}{n}\sum_{i=1}^{n} \hat{TP}^{i}_{j}(f) \qquad \hat{\theta}^{j}_{macro} = \frac{\sum_{i=1}^{n} \mathbb{I}\{y^{i}_{j} = 0\}}{\sum_{i=1}^{n} \mathbb{I}\{y^{i}_{j} = 1\}}$$

$$\hat{TN}^{j}_{macro}(f) = \frac{1}{n}\sum_{i=1}^{n} \hat{TN}^{i}_{j}(f)$$

$$\hat{F}^{macro}_{\beta}(f) = \frac{1}{L}\sum_{j=1}^{L} \hat{F}^{j}_{\beta}(f^{j})$$

$$\hat{F}^{j}_{\beta}(f^{j}) = \frac{(1+\beta^2)\,\hat{TP}^{j}_{macro}(f^{j})}{\beta^2 + \hat{\theta}^{j}_{macro} + \hat{TP}^{j}_{macro}(f) - \hat{\theta}^{j}_{macro}\cdot\hat{TN}^{j}_{macro}(f)}$$

Similarly as before, $\hat{F}^{j}_{\beta}(f^{j}) \geq v^{j} \Leftrightarrow \underbrace{(1+\beta^2 - v^{j})}\cdot \hat{TP}^{j}_{micro}(f^{j}) + v^{j}\,\hat{\theta}^{j}_{macro}\,\hat{TN}^{j}_{macro}(f) \geq v^{j}(\beta^2 + \hat{\theta}^{j}_{macro})$

AMP_{macro}

$$v^{j}_{t+1} \leftarrow \hat{F}^{j}_{\beta}(f^{j}_{t})$$

$$f^{j}_{t+1} \leftarrow \arg\max_{f^{j}} \left\{ a^{j}_{t+1}\,\hat{TP}^{j}_{macro}(f^{j}) + b^{j}_{t+1}\,\hat{TN}^{j}_{macro}(f^{j}) \right\}$$

$$\boxed{\begin{aligned} a^{j}_{t} &= (1+\beta^2 - v^{j}_{t}) \\ b^{j}_{t} &= v^{j}_{t}\cdot\hat{\theta}^{j}_{macro} \end{aligned}}$$

$$f_{t+1} = \{f^{1}_{t+1}, f^{2}_{t+1}, \dots f^{L}_{t+1}\}$$

$\therefore$ f^{j}_{t+1} is solved by obtaining a cost weighted classification problem with the costs a^{j}_{t} and b^{j}_{t} on the positives (points that have the label j) and negatives (points which do not have label j)

Minimizing the instance weighted F-measure on data is an NP-hard problem in general.